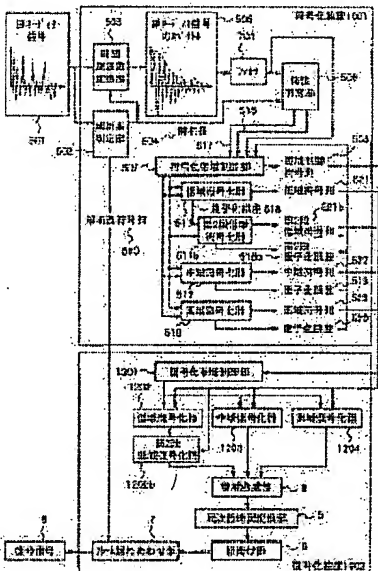


# AUDIO SIGNAL ENCODING METHOD AND AUDIO SIGNAL DECODING METHOD

**Patent number:** JP2003058196 (A)  
**Publication date:** 2003-02-28  
**Inventor(s):** TSUSHIMA MINEO; NORIMATSU TAKESHI; ISHIKAWA TOMOKAZU  
**Applicant(s):** MATSUSHITA ELECTRIC IND CO LTD  
**Classification:**  
 - international: G10L19/02; G10L19/00; H03M7/30; G10L19/00; H03M7/30; (IPC1-7): G10L19/02; G10L19/00; H03M7/30  
 - european:  
**Application number:** JP20020211570 20020719  
**Priority number(s):** JP20020211570 20020719; JP19980059403 19980311

## Abstract of JP 2003058196 (A)

**PROBLEM TO BE SOLVED:** To provide an audio signal encoding method and an audio signal decoding method for adaptively performing scalable coding so that sufficient performance can be presented when encoding various kinds of audio signals. **SOLUTION:** Adaptive scalable coding is performed for changing the frequency range of encoding by automatically deciding the character/distribution of source audio signals without performing fixed scalable coding.



Data supplied from the esp@cenet database — Worldwide

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-58196

(P2003-58196A)

(43) 公開日 平成15年2月28日 (2003.2.28)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テマコード* (参考)
G 1 0 L 19/02		H 0 3 M 7/30	A 5 D 0 4 5
19/00		G 1 0 L 7/04	G 5 J 0 6 4
H 0 3 M 7/30		9/18	M

審査請求 未請求 請求項の数 4 O L (全 19 頁)

(21) 出願番号 特願2002-211570(P2002-211570)  
(62) 分割の表示 特願平11-65335の分割  
(22) 出願日 平成11年3月11日(1999.3.11)  
(31) 優先権主張番号 特願平10-59403  
(32) 優先日 平成10年3月11日(1998.3.11)  
(33) 優先権主張国 日本 (J P)

(71) 出願人 000005821  
松下電器産業株式会社  
大阪府門真市大字門真1006番地  
(72) 発明者 津島 峰生  
大阪府門真市大字門真1006番地 松下電器  
産業株式会社内  
(72) 発明者 則松 武志  
大阪府門真市大字門真1006番地 松下電器  
産業株式会社内  
(74) 代理人 100081813  
弁理士 早瀬 憲一

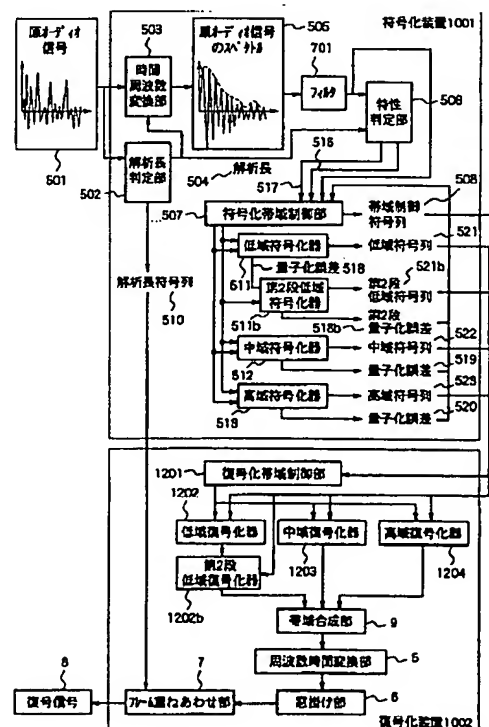
最終頁に続く

(54) 【発明の名称】 オーディオ信号符号化方法、及びオーディオ信号復号化方法

(57) 【要約】

【課題】 従来方式の固定スケーラブルコーディングでは、低域、中域、高域の各帯域量子化器が固定されているため、原オーディオ信号の分布により量子化誤差をできるだけ少なくするように符号化することが困難であった。それゆえ、多種、多様な性質や分布を持つオーディオ信号の符号化を行う際には、十分な性能を発揮できず、高音質で、高効率なスケーラブルコーディングを行なうことが困難であった。そこで、多種多様なオーディオ信号の符号化を行なう際に、十分な性能を発揮するように適応的にスケーラブルコーディングするオーディオ信号符号化、復号化方法を提供する。

【解決手段】 固定スケーラブルコーディングを行わず、原オーディオ信号の性質・分布を自動的に判定し、符号化する周波数範囲を変化させる適応スケーラブルコーディングを行うようにした。



## 【特許請求の範囲】

【請求項 1】 特性判定ステップ、符号化帯域制御ステップ、符号化ステップを包含し、時間一周波数変換されたオーディオ信号を符号化列に変換するオーディオ信号符号化方法であって、

符号化列は、符号化情報と帯域制御符号列とを含み、符号化ステップは、複数の符号化サブステップを有し、符号化帯域制御ステップの制御によりオーディオ信号の多段符号化を行い符号化情報を出力し、

特性判定ステップは、入力されるオーディオ信号を判定し、符号化する各周波数帯域の重み付けを示す帯域重み情報を出力し、

符号化帯域制御ステップは、帯域重み情報に基づいて、多段符号化を構成する各符号化サブステップの量子化帯域、接続順を決定し、決定した各符号化サブステップの量子化帯域、接続順に基づいてスケラブルに構成される多段符号化を符号化ステップに行わせ、

決定した各符号化サブステップの量子化帯域、接続順を示す帯域制御符号列を出力する、

オーディオ信号符号化方法。

【請求項 2】 符号化帯域制御ステップは、予め定義された多段符号化のいずれかになるように、各符号化サブステップの量子化帯域、接続順を決定する、請求項 1 記載のオーディオ信号符号化方法。

【請求項 3】 符号化ステップは、量子化誤差を出力し、

符号化帯域制御ステップは、帯域重み情報と量子化誤差とに基づいて、各符号化サブステップの量子化帯域、接続順を決定する、

請求項 1 記載のオーディオ信号符号化方法。

【請求項 4】 復号化帯域制御ステップ、復号化ステップを包含し、符号化情報と帯域制御符号列とを含む符号化列をオーディオ信号に復号するオーディオ信号復号化方法であって、

帯域制御符号列は、符号化情報を多段符号化した際の各符号化の量子化帯域、接続順を示し、

復号化ステップは、複数の復号化サブステップを有し、復号化帯域制御ステップの制御により符号化情報の多段復号化を行い、

復号化帯域制御ステップは、帯域制御符号列に基づいてスケラブルに構成される多段復号化を復号化ステップに行わせる、

オーディオ信号復号化方法。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、オーディオ信号符号化方法、及びオーディオ信号復号化方法に関し、特に、音声信号や音楽信号などのオーディオ信号から得られる特徴量、特にオーディオ信号を直交変換等の手法を

用いて、時間領域から周波数領域に変換した信号を用い、その変換した信号を原オーディオ信号と比較して、できるだけ少ない符号列で表現するために効率的に符号化する方法と、符号化された信号である符号化列のすべて、あるいはその一部のみを用いて、高い品質と広帯域なオーディオ信号を、復号可能な構成の復号化方法に関するものである。

## 【0002】

【従来の技術】オーディオ信号を効率的に符号化、および復号化する様々な手法が提案されている。音楽信号など、20 kHz以上の周波数帯域を有するオーディオ信号の圧縮符号化方式には、MPEGオーディオ方式や、Twin VQ (TC-WVQ) 方式などがある。MPEG方式に代表される符号化方式は、時間軸のデジタルオーディオ信号を、コサイン変換などの直交変換を用いて、周波数軸上のデータに変換し、その周波数軸上の情報を、人間の聴覚的な感度特性を利用して、聴覚的に重要な情報から符号化していく方式であり、聴覚的に重要でない情報や、冗長な情報は符号化しない方式である。一方、Twin VQ (TC-WVQ) 方式は、ベクトル量子化手法を用いて、原デジタル信号の情報量に対して、かなり少ない情報量で表現しようとする符号化方式がある。MPEGオーディオ、および Twin VQ (TC-WVQ) は、それぞれISO/IEC 標準IS-11172-3、およびT.Moriya, H.Suga: An 8 Kbits transform coder for noisy channels, Proc. ICASSP 89, pp196-199、などに述べられている。

【0003】ここで、図10を用いて、一般の、Twin VQ 方式の概要を説明する。原オーディオ信号101を解析長判定部102に入力し、解析長を算出する。また同時に、解析長判定部102は解析長112を量子化し、解析長符号列111を出力する。次に、その解析長112に従って、時間周波数変換部103で、原オーディオ信号101を周波数領域の原オーディオ信号104に変換する。次に、周波数領域の原オーディオ信号104は、正規化処理部（平坦化処理部）106で正規化処理（平坦化処理）され、正規化処理後のオーディオ信号108を得る。正規化処理は、原オーディオ信号104から周波数概形105を計算し、原オーディオ信号104を算出した周波数概形105で割ることにより行われる。さらに、正規化処理部106は、正規化処理に用いた周波数概形情報を量子化し、正規化符号列107を出力する。次に、正規化処理後のオーディオ信号108を、ベクトル量子化部109により量子化し、符号列110が得られる。

【0004】近年、復号器に入力する符号列の一部を用いても、オーディオ信号を再生することができる構造を持つものがある。上記の構造を、スケラブル構造と呼び、スケラブル構造を実現できるように符号化することを、スケラブルコーディングと呼ぶ。

【0005】図11に一般の、Twin VQ 方式で採用され

ている、固定スケーラブルコーディングの一例を示す。原オーディオ信号1301から解析長判定部1303により判定された解析長1314に従って、時間周波数変換部1302により、周波数領域の原オーディオ信号1304を得る。次に、周波数領域の原オーディオ信号1304を、低域符号化器1305に入力すると、量子化誤差1306と、低域符号列1311とが出力される。さらに、量子化誤差1306を中域符号化器1307に入力すると、量子化誤差1308と、中域符号列1312とが出力される。さらに、量子化誤差1308を高域符号化器1309に入力すると、量子化誤差1310と、高域符号列1313とが出力される。ここで、上記低域、または中域、または高域符号化器は、正規化処理部と、ベクトル量子化部とを併せ持ち、その出力は、量子化誤差、および正規化処理部、ならびにベクトル量子化部により出力された各符号列を含む、低域、中域、または高域符号列を、出力するものである。

#### 【0006】

【発明が解決しようとする課題】従来方式の固定スケーラブルコーディングでは、図11に示すように、低域、中域、高域の各帯域量子化器が固定されているため、図12に示すように、原オーディオ信号の分布に対して、量子化誤差をできるだけ少なくするように符号化することが困難であった。それゆえ、多種多様な性質や分布を持つオーディオ信号の符号化を行う際には、十分な性能を発揮できず、高音質で高効率なスケーラブルコーディングを行なうことが困難であった。

【0007】本発明は上記の問題点を解消するためになされたもので、多種多様なオーディオ信号の符号化に際して、オーディオ信号を符号化する際、図13に示すように、多種多様なオーディオ信号を適応的にスケーラブルコーディングすることにより、効率よく、低ビットレートで、かつ、高音質に、符号化を行なうことのできるオーディオ信号符号化方法、及びオーディオ信号復号化方法を提供することを目的としている。

#### 【0008】

【課題を解決するための手段】この課題を解決するために、本発明にかかるオーディオ信号符号化方法、及びオーディオ信号復号化方法は、固定スケーラブルコーディングを用いず、原オーディオ信号の性質、分布にあわせて符号化する周波数範囲を変化させる適応スケーラブルコーディングを行なうようにしたものである。

【0009】本発明に係るオーディオ信号符号化方法は、特性判定ステップ、符号化帯域制御ステップ、符号化ステップを包含し、時間一周波数変換されたオーディオ信号を符号化列に変換するオーディオ信号符号化方法であって、符号化列は、符号化情報と帯域制御符号列とを含み、符号化ステップは、複数の符号化サブステップを有し、符号化帯域制御ステップの制御によりオーディオ信号の多段符号化を行い符号化情報を出力し、特性判

定ステップは、入力されるオーディオ信号を判定し、符号化する各周波数帯域の重み付けを示す帯域重み情報を出力し、符号化帯域制御ステップは、帯域重み情報に基づいて、多段符号化を構成する各符号化サブステップの量子化帯域、接続順を決定し、決定した各符号化サブステップの量子化帯域、接続順に基づいてスケーラブルに構成される多段符号化を符号化ステップに行わせ、決定した各符号化サブステップの量子化帯域、接続順を示す帯域制御符号列を出力するものである。

【0010】本発明に係るオーディオ信号符号化方法は、前記オーディオ信号符号化方法において、符号化帯域制御ステップが、予め定義された多段符号化のいずれかになるように、各符号化サブステップの量子化帯域、接続順を決定するようにしたものである。

【0011】本発明に係るオーディオ信号符号化方法は、前記オーディオ信号符号化方法において、符号化ステップが、量子化誤差を出力し、符号化帯域制御ステップが、帯域重み情報と量子化誤差とに基づいて、各符号化サブステップの量子化帯域、接続順を決定するようにしたものである。

【0012】本発明に係るオーディオ信号復号化方法は、復号化帯域制御ステップ、復号化ステップを包含し、符号化情報と帯域制御符号列とを含む符号化列をオーディオ信号に復号するオーディオ信号復号化方法であって、帯域制御符号列は、符号化情報を多段符号化した際の各符号化の量子化帯域、接続順を示し、復号化ステップは、複数の復号化サブステップを有し、復号化帯域制御ステップの制御により符号化情報の多段復号化を行い、復号化帯域制御ステップは、帯域制御符号列に基づいてスケーラブルに構成される多段復号化を復号化ステップに行わせるようにしたものである。

#### 【0013】

【発明の実施の形態】以下、本発明の実施の形態1について、図1ないし図9を用いて、また、実施の形態2について、図14ないし図20を用いて、説明する。

【0014】（実施の形態1）図1は、本発明の実施の形態1による、適応スケーラブルコーディングを行なう、オーディオ信号符号化装置のブロック図を示す。図1において、1001は原オーディオ信号501を符号化する符号化装置である。該符号化装置1001において、502は上記原オーディオ信号501を解析する際の解析長504を判定する解析長判定部、503は上記解析長504の単位で、原オーディオ信号501の時間軸を周波数軸に変換する時間周波数変換部、504は上記解析長判定部502で判定された解析長、505は原オーディオ信号のスペクトル、701は該原オーディオ信号のスペクトル505が入力されるフィルタ、506は原オーディオ信号のスペクトル505の特性を判定し、上記符号化装置1001における複数の各段の各符号化器511、512、513、511b等、の量子化

するオーディオ信号の周波数帯域を決定する特性判定部、507は該特性判定部506で決定された各符号化器の周波数帯域と、上記周波数変換されたオーディオ信号をその入力とし、複数の各段の各符号化器512, 513, 514, 511b等、の接続順を決定し、各符号化器の量子化帯域、及び接続順を符号列に変換する符号化帯域制御部、508は、該符号化帯域制御部507より出力される上記符号列である帯域制御符号列、510は上記解析長判定部502より出力された上記解析長504を符号列とした解析長符号列、511, 512, 513は、上述した、それぞれ低域、中域、高域の信号を符号化する低域符号化器、中域符号化器、高域符号化器、511bは第1段の低域符号化器511の量子化誤差518を符号化する第2段低域符号化器、521, 522, 523は該各符号化器511, 512, 513から出力される符号化信号である低域符号列、中域符号列、高域符号列、521bは第2段低域符号化器511bの符号化出力である第2段低域符号列、518, 519, 520は該各符号化器511, 512, 513から出力される、符号化される前の信号と上記各符号化信号との差である量子化誤差、518bは第2段低域符号化器511bの量子化誤差である第2段量子化誤差である。

【0015】一方、1002は上記符号化装置1001で符号化された符号化列を復号化する復号化装置である。該復号化装置1002において、5は上記符号化装置1001における時間周波数変換部503と逆の変換を行なう周波数時間変換部、6は時間軸上で窓関数を乗じる窓掛けを行なう窓掛け部、7はフレーム重ね合わせ部、8は復号信号、9は帯域合成部、1201は復号化帯域制御部、1202, 1203, 1204は、それぞれ上記低域符号化器、中域符号化器、高域符号化器511, 512, 513に対応して、復号化を行なう低域復号化器、中域復号化器、高域復号化器、1202bは第1段低域復号化器1202の出力を復号化する第2段低域復号化器である。

【0016】ここで、第2段以降の符号化器、復号化器はさらに他の帯域にも、またさらに、多段にも設けてもよいものであり、これが多段になるほど、必要に応じて、符号化、復号化の精度を向上できるものである。

【0017】以下、先ず、符号化装置1001の動作について説明する。符号化しようとする原オーディオ信号501は、時間的に連続するデジタル信号系列であるとする。例えば、音声信号を、サンプリング周波数48kHzで16ビットに量子化したデジタル信号であるとする。

【0018】上記原オーディオ信号501を解析長判定部502に入力する。上記解析長判定部502は、入力された上記原オーディオ信号501の特性を判断し、解析長504を決定し、その結果は解析長符号列510と

して、復号化装置1002に送られる。解析長504としては、たとえば256、1024、4096などが用いられる。例えば、原オーディオ信号501に含まれる高域周波数成分が所定の値を超える場合には、解析長504を256とし、低域周波数成分が所定の値を超え、かつ高域周波数成分が所定の値より小さい場合には、解析長504を4096とし、それ以外の場合は、解析長504を1024とする。こうして決定された解析長504に従って、時間周波数変換部503により原オーディオ信号501のスペクトル505を算出する。

【0019】図2に、本発明の実施の形態1によるオーディオ信号符号化装置における、時間周波数変換部503のブロック図を示す。上記原オーディオ信号501は、そのサンプル値が所定のサンプル数に達するまでフレーム分割部201で蓄積され、該蓄積されたサンプル数が、上記解析長判定部502で決定された解析長504に達すると、出力を行なう。また、フレーム分割部201は、あるシフト長ごとに出力を行う構成のものであり、例えば、解析長504を4096サンプルとした場合において、解析長504の半分のシフト長を設定すれば、解析長504が2048サンプルに達するに相当する時間ごとに、最新の4096サンプルを出力するなどの構成を持つ。当然ながら、解析長504や、サンプリング周波数が変わっても、同様に、シフト長を解析長504の半分に設定した構成を持つことは可能である。そして、このフレーム分割部201からの出力は、後段の窓掛け部202へと入力される。窓掛け部202では、フレーム分割部201からの出力に対して、時間軸上で窓関数を乗じて、窓掛け部202の出力とする。この様子は、例えば、(数1)で示される。

【0020】

【数1】

$$h_{xi} = h_i * x_i \quad i = 1, 2, \dots, N$$

$$h_i = \sin \left\{ \frac{\pi}{N} (i + 0.5) \right\}$$

ただし、ここで、 $x_i$ はフレーム分割部201からの出力で、 $h_i$ は窓関数、 $h_{xi}$ は窓掛け部202からの出力である。まだ、 $i$ は時間のサフィックスである。なお、(数1)で示した窓関数 $h_i$ は一例であり、窓関数は必ずしも、(数1)のものである必要はない。

【0021】窓関数の選択は、窓掛け部202に入力される信号の特徴と、フレーム分割部201の解析長504と、時間的に前後に位置するフレームにおける窓関数の形状とに依存する。例えば、窓掛け部202に入力される信号の特徴として、フレーム分割部201の解析長504を $N$ とした場合、 $N/4$ ごとに入力される信号の平均パワーを算出して、その平均パワーが非常に大きく変動する場合は、解析長504を $N$ よりも短くして(数1)に示した演算を実行する、などの選択を行う。ま

た、前の時刻のフレームの窓関数の形状と、後ろのフレームの窓関数の形状とに応じて、現在の時刻のフレームの窓関数の形状に歪みがないように、適宜選択するのが望ましい。

【0022】次いで、窓掛け部202からの出力は、MDCT部203に入力され、ここで変形離散コサイン変換が施され、MDCT係数が出力される。変形離散コサイン変換の一般式は、(数2)で表される。

【0023】

【数2】

$$y_k = \sum_{n=0}^{N-1} h_{x_n} * \cos \left\{ \frac{2\pi \left( k + \frac{1}{2} \right) (n + n_0)}{N} \right\}$$

$$n_0 = \frac{N}{4} + \frac{1}{2} \left( k = 0, 1, \dots, \frac{N}{2} - 1 \right)$$

このようにMDCT部203の出力であるMDCT係数は、(数2)中の、 $y_k$ で表せるとすると、MDCT部203の出力は周波数特性を示し、 $y_k$ の変数 $k$ が0に近いほど、低い周波数成分に、0から増大して $N/2-1$ に近くなるほど、高い周波数成分に、線形に対応する。こうして算出された上記MDCT係数が、原オーディオ信号のスペクトル505となる。

【0024】次に、上記原オーディオ信号のスペクトル505をフィルタ701へと入力する。該フィルタ701の入力を、 $x_{701}(i)$ 、出力を、 $y_{701}(i)$ とすると、例えば、(数3)で表されるフィルタを用いる。

【0025】

【数3】

$$y_{701}(i) = w_{701}(i) * \{x_{701}(i) + x_{701}(i+1)\}$$

$$i = 0, 1, \dots, fs-2$$

ここで、 $fs$ は解析長504である。(数3)で表されるフィルタ701は、一種の移動平均フィルタであるが、当然ながら、移動平均フィルタに限定する必要はなく、他の、たとえば高域通過フィルタであってもいいし、帯域抑制フィルタであっても良い。

【0026】フィルタ701の出力と、解析長判定部502で算出した解析長504とを、特性判定部506に入力する。図6に、特性判定部506の詳細を示す。特性判定部506では、原オーディオ信号501、および原オーディオ信号のスペクトル505、の聴覚的、物理的な特性を決定する。原オーディオ信号501、および該スペクトル505の聴覚的、物理的特性とは、例えば、音声か、音楽か、の違いである。音声の場合、たとえば6kHzより低域に、大半の周波数成分があるものである。

【0027】次に、特性判定部506の動作を、図6を用いて説明する。特性判定部506に入力された原オーディオ信号のスペクトル505をフィルタ701によ

てフィルタリングした信号を、 $x_{506}(i)$ とすると、この $x_{506}(i)$ を基に、スペクトルパワー $p_{506}(i)$ を、(数4)により、スペクトルパワー計算部803で計算する。

【0028】

【数4】

このスペクトルパワー $p_{506}(i) = x_{506}(i)^2$ を、符号化帯域制御部507の入力の一つとし、各符号化器の帯域制御重み517とする。また、解析長504が小さい場合、例えば256なるとき、各符号化器を固定的に配置するよう、配置決定部804で決定し、符号化帯域制御部507へと、符号化帯域配置情報516を、固定配置として送る。

【0029】解析長504が小さい場合以外の場合、たとえば4096や1024のときは、各符号化器を動的に配置するよう、配置決定部804で決定し、符号化帯域制御部507へと、符号化帯域配置情報516を、動的配置として送る。

【0030】次に、符号化帯域制御部507の動作を、図7を用いて説明する。符号化帯域制御部507には、上記特性判定部506からの出力である帯域制御重み517と、符号化帯域配置情報516、および原オーディオ信号のスペクトル505をフィルタ701でフィルタリングした信号と、各符号化器の出力した量子化誤差518、または519、または520が入力される。ただし、これらの入力があるのは、各符号化器511、512、513、511bと、符号化帯域制御部507とが、再帰的に動作するためであり、初回の符号化帯域制御部507の動作においては、量子化誤差がないため、量子化誤差を除いた3つの入力となる。

【0031】上記のように、解析長504が小さく、符号化帯域配置情報516が固定配置となる場合は、予め定義された帯域の固定配置に従って、符号化を、低域から中域、高域へと順に実行するよう、量子化順序決定部902、および、符号化器数決定部903、帯域幅算出部901により、符号化器の量子化帯域、個数、接続順を決定し、符号化を行う。即ち、その時の帯域制御符号列508には、符号化器の帯域情報、符号化器数、および、その接続順序が、情報として符号化される。

【0032】たとえば、各符号化器の符号化帯域、および符号化器数を、それぞれ0Hz～4kHzに1つ、0Hz～8kHzに1つ、4kHz～12kHzに1つ、8kHz～16kHzに2つ、16kHz～24kHzに3つ、となるように、符号化器を配置し、符号化を行う。

【0033】次に、符号化帯域配置情報516が動的配置になっている場合の、符号化帯域制御部507の動作について説明する。符号化帯域制御部507は、各符号化器の量子化帯域幅を決定する帯域幅算出部901、各符号化器の量子化順序を決定する量子化順序決定部90

2、さらに各帯域の符号化器の数を決定する符号化器数決定部903、の3つよりなる。符号化帯域制御部507に入力された信号をもとに、各符号化器の帯域幅を決定する訳であるが、所定の帯域、例えば、0Hz～4kHz、0kHz～8kHz、4kHz～12kHz、8kHz～16kHz、16kHz～24kHzの各帯域において、帯域制御重み517、および各

符号化器が符号化した後の量子化誤差、を乗算したものの平均値を算出する。ここで、帯域制御重み517を、weight517(i)、量子化誤差を、err507(i)とすると、  
(数5)により、平均値を算出する。

【0034】

【数5】

$$\text{Ave}_{901}(j) = \frac{1}{f_{\text{upper}}(j) - f_{\text{lower}}(j)} \sum_{i=f_{\text{upper}}(j)}^{f_{\text{lower}}(j)} \text{weight}_{517}(i) * \text{err}_{507}(i)^2$$

ここで、jは各帯域のインデックス、Ave901(j)は、帯域jにおける平均値、fupper(j)、およびflower(j)は、帯域jの上限周波数、および下限周波数である。こうして得られた平均値 Ave901(j)が最大となるjを検索し、それが、符号化器が符号化する帯域となる。さらに、検索されたjの値を、符号化器数決定部903に送り、jに対応する帯域の符号化器数を一つ増やすようにし、所定の符号化帯域にいくつの符号化器が存在するかを記憶しておき、記憶している符号化器数の合計が、予め決定しておいた符号化器の総数になるまで、符号化を繰り返す。最後に、符号化器の帯域、および符号化器数を、帯域制御符号列508として、復号化器へと伝送する。

【0035】次に、符号化器3の動作について、図3を用いて説明する。符号化器3は、正規化部301と、量子化部302とからなる。正規化部301では、フレーム分割部201からの出力である時間軸の信号と、MDCT部203からの出力であるMDCT係数、との両者を入力として、いくつかのパラメータを用いて、MDCT係数を正規化する。ここで、MDCT係数の正規化とは、低域成分と高域成分とで非常に大きさに違いのあるMDCT係数の大きさのばらつきを抑圧することを意味し、例えば、低域成分が高域成分に対して非常に大きい場合などは、低域成分では大きな値、高域成分では小さな値、となるようなパラメータを選出し、これで上記MDCT係数を除算することにより、MDCT係数の大きさのばらつきを抑圧することを指す。また正規化部301では、正規化に用いたパラメータを表現するインデックスを、正規化符号列303として符号化する。

【0036】量子化部302では、正規化部301で正

規化されたMDCT係数を入力として、MDCT係数の量子化を行う。この際、該量子化部302は、該量子化した値と、コードブック中にある複数のコードインデックスに対応する各量子化出力、との間の差が最も小さくなるような、そのような該コードインデックスを出力する。この場合、上記量子化部302で量子化した値と、該量子化部302から出力されるコードインデックスに対応する値、との差が量子化誤差である。

【0037】次に、図4を用いて、上記正規化部301の詳細な一例を説明する。図4において、401はフレーム分割部201とMDCT部203の出力を受ける周波数概形正規化部、402は上記周波数概形正規化部401の出力を受け、帯域テーブル403を参照して、正規化を行う帯域振幅正規化部である。

【0038】次に、動作について説明する。周波数概形正規化部401では、フレーム分割部201からの時間軸上のデータ出力を用いて、大まかな周波数の概形である周波数概形を算出し、MDCT部203からの出力であるMDCT係数を除算する。周波数概形を表現するのに用いたパラメータは、正規化符号列303として符号化される。帯域振幅正規化部402では、周波数概形正規化部401からの出力信号を入力として、帯域テーブル403で示された帯域ごとに正規化を行う。例えば、周波数概形正規化部401の出力であるMDCT係数が、dct(i)(i=0～2047)とし、帯域テーブル403が、例えば、(表1)に示されるようなものであるとすると、(数6)などを用いて、各帯域毎の振幅の平均値を算出する。

【0039】

【表1】

帯域 k	f <sub>lower</sub> (k)	f <sub>upper</sub> (k)
0	0	10
1	11	22
2	23	33
3	34	45
4	46	56
5	57	68
6	69	80
7	81	92
8	93	104
9	105	116
10	117	128
11	129	141
12	142	153
13	154	166
14	167	179
15	180	192
16	193	205
17	206	219
18	220	233
19	234	247
20	248	261
21	262	276
22	277	291
23	292	307
24	308	323
25	324	339
26	340	356
27	357	374
28	375	392
29	393	410
30	411	430
31	431	450
32	451	470
33	471	492
34	493	515
35	516	538
36	539	563
37	564	587
38	589	615
39	616	643
40	645	673
41	674	705
42	706	737
43	738	772
44	773	809
45	810	848
46	849	889
47	890	932
48	933	978
49	979	1027
50	1028	1079

帯域 k	f <sub>lower</sub> (k)	f <sub>upper</sub> (k)
51	1080	1135
52	1136	1193
53	1194	1255
54	1256	1320
55	1321	1389
56	1390	1462
57	1463	1538
58	1539	1617
59	1618	1699
60	1700	1783
61	1784	1870
62	1871	1958
63	1959	2048

【数6】

$$\text{sum}_j = \sum_{i=\text{bjlow}}^{\text{bjhigh}} \text{dct}(i)^p$$

$$\text{ave}_j = \left( \frac{\text{sum}_j}{\text{bjhigh} - \text{bjlow} + 1} \right)^{-p} \quad \text{bjlow} \leq i \leq \text{bjhigh}$$

ここで、bjlow, bjhighは、帯域テーブル403に示されたj番目の帯域におけるdct(i)が属する最も低域のイン

$$n\_dct(i) = \text{dct}(i) / \text{ave}_j$$

ave<sub>j</sub>の量子化は、スカラーの量子化を用いてもよいし、コードブックを用いてベクトル量子化を行ってもよい。帯域振幅正規化部402では、qave<sub>j</sub>を表現するのに用いたパラメータのインデックスを、正規化符号列303として符号化する。

【0041】なお、符号化器における正規化部301の

デックスiと、最も高域のインデックスiをそれぞれ示している。また、pは距離計算におけるノルムであり、2などが望ましい。ave<sub>j</sub>は、各帯域番号jにおける振幅の平均値である。帯域振幅正規化部402では、ave<sub>j</sub>を量子化して、qave<sub>j</sub>を算出して、例えば、(数7)を用いて正規化する。

【0040】

【数7】

$$\text{bjlow} \leq i \leq \text{bjhigh}$$

構成は、図4の周波数概形正規化部401と、帯域振幅正規化部402、との両者を用いた構成のものを示したが、周波数概形正規化部401のみを用いた構成でもよく、帯域振幅正規化部402のみを用いた構成でもよい。さらに、MDCT部203から出力されるMDCT係数の低域成分と、高域成分とで大きなばらつきがない



場合には、上記両者を用いない構成で、MDC T部203の出力信号を、そのまま量子化部302に inputsする構成としてもよい。

【0042】次に、図5を用いて、図4の周波数概形正規化部401の詳細について説明する。図5において、601はフレーム分割部201の出力を受ける線形予測分析部、602は線形予測分析部601の出力を受ける概形量子化部、603はMDC T部203の出力を受ける包絡特性正規化部である。

【0043】次に、上記周波数概形正規化部401の動作について、図5を参照して説明する。上記線形予測分析部601では、フレーム分割部201からの時間軸上のオーディオ信号を入力として、線形予測分析(Linear Predictive Coding)を行う。線形予測分析の線形予測係数(LPC係数)は、ハミング窓などの窓掛けされた信号の自己相関関数を算出し、正規方程式などを解くことで、一般に算出可能である。算出された線形予測係数は、線スペクトル対係数(LSP(Line Spectrum Pair)係数)などに変換され、概形量子化部602で量子化される。ここでの量子化手法としては、ベクトル量子化を用いてもよいし、スカラー量子化を用いてもよい。そして、概形量子化部602で量子化されたパラメータが表現する周波数伝達特性を、包絡特性正規化部603で算出し、MDC T部203からの出力であるMDC T係数を、これで除算することによって正規化する。具体的な算出例としては、概形量子化部602で量子化されたパラメータと等価な線形予測係数を、 $qlpc(i)$ とすれば、包絡特性正規化部603で算出される上記周波数伝達特性は、例えば、(数8)で表すことができる。

【0044】

【数8】

$$li = \begin{cases} qlpc(i) & 0 \leq i \leq ORDER \\ 0 & ORDER+1 \leq i \leq N \end{cases}$$

$$env(i) = \frac{1}{fft(li)}$$

$$subvector_i(j) = \begin{cases} vector \left( \frac{VTOTAL}{CR} \cdot i+j \right) & \text{ただし } \frac{VTOTAL}{CR} \cdot i+j < TOTAL \\ 0 & \text{ただし } \frac{VTOTAL}{CR} \cdot i+j \geq TOTAL \end{cases}$$

ここで、 $i$  番目の音源サブベクトルの  $j$  番目の要素は  $subvector_i(j)$  であり、MDC T係数1001は  $vector()$  であり、MDC T係数1001の総要素数がTOTALで、音源サブベクトル1003の要素数がCR、VTOTALは、TOTALと同じ値かより大きい値で、VTOTAL/CRが正数値になるように設定する。例えば、TOTALが2048の時、CRが19で、VTOTALが2052、CRが23で、VTOTALが2070、CRが21で、VTOTALが2079などである。重みサブベクトル1004も、数10の

ここで、ORDERは10~40くらいが望ましい。 $fft()$ は高速フーリエ変換を意味する。算出された周波数伝達特性 $env(i)$ を用いて、包絡特性正規化部603では、例えば、下記に示す(数9)を用いて、正規化を行う。

【数9】

$$fdct(i) = \frac{mdct(i)}{env(i)}$$

ここで、 $mdct(i)$ はMDC T部203からの出力信号で、 $fdct(i)$ は正規化された包絡特性正規化部603からの出力信号である。

【0045】次に図8を用いて、上記符号化装置1における量子化部302の量子化方法の詳細な動作について説明する。量子化部302に inputsされるMDC T係数1001は、そのMDC T係数1001から幾つかを抜き出して、音源サブベクトル1003を構成する。同様に、正規化部301で、正規化部301の入力であるMDC T係数を、正規化部301の出力であるMDC T係数で割った係数列を、正規化成分1002とした時、この正規化成分1002についても、MDC T係数1001から音源サブベクトル1003を抜き出したのと同じ規則で、該正規化成分1002からサブベクトルの抽出を行い、重みサブベクトル1004を構成することができる。音源サブベクトル1003、および重みサブベクトル1004を、MDC T係数1001および正規化成分1002からそれぞれ抽出する規則は、例えば、(数10)で示す方法などがある。

【0046】

【数10】

手順で抽出可能である。ベクトル量子化器1005では、コードブック1009中のコードベクトルの中から、音源サブベクトル1003との距離が、重みサブベクトル1004で重み付けて最も小さくなるものを探し、その最小の距離を与えたコードベクトルのインデックスと、最小の距離を与えたコードベクトルと入力音源サブベクトル1003との量子化誤差に相当する残差サブベクトル1010とを出力する。

【0047】実際の計算手順例においては、ベクトル量

子化器1005が、距離計算手段1006、コード決定手段1007、残差生成手段1008、の3つの構成要素からなるものとして、説明する。距離計算手段1006では、例えば、(数11)を用いて、i番目の音源サブベクトル1003と、コードブック1009のk番目のコードベクトル、との距離を算出する。

【0048】

【数11】

$$dik = \sum_{j=0}^{CR-1} w_j^R (\text{subvector}_i(j) - C_k(j))^S$$

ここで、 $w_j$ は、重みサブベクトルのj番目の要素、 $C_k(j)$ は、k番目のコードベクトルのj番目の要素、 $R$ 、 $S$ は、距離計算のノルムであり、 $R$ 、 $S$ の値としては、1、1.5、2などが望ましい。なお、このノルム $R$ と $S$ は、同一の値である必要はない。 $dik$ は、i番目の音源サブベクトルに対するk番目のコードベクトルの距離を意味する。コード決定手段1007では、(数11)などで算出された距離の中で、最小となるコードベクトルを選出し、そのインデックスを符号列304として符号化する。例えば、複数の上記 $dik$ があるうちの、 $d_{iu}$ が最小値である場合、i番目のサブベクトルに対する符号化されるインデックスは、 $u$ となる。残差生成手段1008では、コード決定手段1007で選出したコードベクトルを用いて、(数12)により残差サブベクトル1010を生成する。

【0049】

【数12】

$$\text{res}_i(j) = \text{subvector}_i(j) - C_u(j)$$

ここで、i番目の残差サブベクトル1010のj番目の要素は、 $\text{res}_i(j)$ であり、コード決定手段1007で選出されたコードベクトルのj番目の要素を、 $C_u(j)$ とする。上記残差サブベクトル1010を用いて(数10)の逆過程の演算を行ってベクトルを求め、該ベクトルと、当該符号化器の元々の符号化対象であったベクトルとの差を、それ以降の各符号化器の量子化対象となるMDCT係数として保持する。ただし、ある帯域の符号化が、それ以降の符号化器に影響を与えない帯域に対して符号化を行っている場合、つまり、以降の符号化器が符号化をすることがない場合は、残差生成手段1008による、残差サブベクトル1010、MDCT1011の生成は必要ない。なお、コードブック1009が持つコードベクトルの個数はいくつでもよいが、メモリ容量、計算時間等を考慮すると、64程度とすることが好ましい。

【0050】なお、上記ベクトル量子化器1005の他の例としては、以下のような構成も可能である。すなわち、距離計算手段1006では、(数13)を用いて距離を算出する。

【0051】

【数13】

$$dik = \begin{cases} \sum_{j=0}^{CR-1} w_j^R (\text{subvector}_i(j) - C_k(j))^S & k < K \\ \sum_{j=0}^{CR-1} w_j^R (\text{subvector}_i(j) - C_{K-k}(j))^S & k \geq K \end{cases}$$

ただし、 $K$ は、コードブック1009のコード検索に用いるコードベクトルの総数である。コード決定手段1007では、(数13)で算出された距離 $dik$ の最小値を与える $k$ を選出し、そのインデックスを符号化する。ただし、 $k$ は0から $2K-1$ までの値となる。残差生成手段1008では、(数14)を用いて残差サブベクトル1010を生成する。

【0052】

【数14】

$$\text{res}_i(j) = \begin{cases} \text{subvector}_i(j) - C_u(j) & 0 \leq k < K \\ \text{subvector}_i(j) + C_u(j) & K \leq k < 2K \end{cases}$$

ここで、コードブック1009が持つコードベクトルの個数はいくつでもよいが、メモリの容量、計算時間等を考慮すると、64程度とすることが好ましい。また、上記では、重みサブベクトル1004を、正規化成分1002のみから生成する構成について述べたが、重みサブベクトル1004に、人間の聴覚特性を考慮した重みをさらに乗じて、重みサブベクトルを生成することも可能である。以上のようにして、複数の各段の各符号化器の帯域幅、符号化器の個数、及び、接続順序が動的に決定される。そして、こうして決定された各符号化器の情報を基に、量子化を行なう。

【0053】一方、復号化装置1002では、各帯域の符号化器の出力である正規化符号列と、該正規化符号列に対応した量子化部からの符号列、さらに符号化装置における符号化帯域制御部の出力である帯域制御符号列、また解析長判定部の出力である解析長符号列、を用いて、復号を行う。

【0054】図9に、復号化器1202、1203、…の構成を示す。各復号化器は、正規化されたMDCT係数を再生する逆量子化部1101と、正規化係数を復号し、上記再生された正規化されたMDCT係数と、正規化係数とを乗算する逆正規化部1102とからなる。

【0055】逆正規化部1102では、各符号化器の正規化部301からの正規化符号列303から、符号化装置1で正規化に用いたパラメータの復元を行い、逆量子化部1101の出力と、該パラメータとを乗算し、MDCT係数の復元を行う。

【0056】復号化帯域制御部1201では、符号化帯域制御部507の出力である帯域制御符号列508を用いて、符号化装置で用いた符号化器の配置や、符号化器の個数の情報を復元し、その情報に基づいて各帯域に各

復号化器 1202、1203、1204、1202b を配置し、符号化装置での各符号化器 511、512、513、511b の符号化順序とは逆順に帯域を合成する帯域合成部 9 により、MDCT 係数を得る。こうして得られた該 MDCT 係数を入力とする周波数時間変換部 5 では、逆 MDCT を行い、周波数領域の信号から時間領域の信号への復元を行う。上記逆 MDCT 係数の計算は、例えば、

【0057】

【数 15】

$$xx(n) = \frac{2}{N} \sum_{k=0}^{N-1} yy_k \cos \left\{ \frac{2\pi(k+1/2)(n+n_0)}{N} \right\}$$

$$n_0 = \frac{N}{4} + \frac{1}{2}$$

ここで、 $yy_k$  は帯域合成部 9 で復元された MDCT 係数で、 $xx(n)$  は逆 MDCT 係数であり、これを周波数時間変換部 5 の出力とする。窓掛け部 6 では、周波数時間変換部 5 からの出力  $xx(i)$  を用いて窓掛けを行う。窓掛けは、符号化装置 1 の時間周波数変換部 503 にある窓掛け部 202 で用いた窓を用い、たとえば、

【数 16】で示される処理を行う。

【0058】

【数 16】

$$z(i) = xx(i) * h_i$$

ここで、 $z(i)$  は窓掛け部 6 の出力である。フレーム重ね合わせ部 7 では、窓掛け部 6 からの出力を用いて、オーディオ信号を再生する。窓掛け部 6 からの出力は、時間的に重複した信号となっているので、フレーム重ね合わせ部 7 では、例えば、【数 17】を用いて、復号化装置 1002 の出力信号とする。

【0059】

【数 17】

$$out_m(i) = z_m(i) + z_{m-1}(i + \text{SHIFT})$$

ここで、 $z_m(i)$  は、第  $m$  時刻フレームの第  $i$  番目の窓掛け部 6 の出力信号  $z(i)$  で、 $z_{m-1}(i)$  は、第  $m-1$  時刻フレームの第  $i$  番目の窓掛け部 6 の出力信号とし、SHIFT は、符号化装置の解析長 504 に相当するサンプル数、 $out_m(i)$  は、フレーム重ね合わせ部 7 の第  $m$  時刻フレームにおける復号化装置 1002 の出力信号とする。

【0060】また、本実施の形態 1 においては、以下のように、符号化帯域制御部 507 において、帯域幅算出部 901 で算出する量子化可能周波数範囲を、解析長 504 により制限することもある。たとえば、解析長 504 が 256 の場合、各符号化器の量子化可能周波数範囲の下限を 4 kHz 程度、上限を 24 kHz 程度にする。解析長が 1024 あるいは 2048 の場合、下限を 0 Hz、上限を 16 kHz 程度にする。さらに一度、解析長 504 が 256 になれば、その後一定時間の間、たとえば 20 m

sec 程度の間、各量子化器の量子化可能周波数範囲や、量子化器の配置を固定するように、量子化順序決定部 902 により制御することもできる。この処理を用いることにより、経時的に量子化器の配置を一定にし、聴感的な帯域の出入り感（ある瞬間まで高い帯域が中心の音声であったものが、突然、低い帯域が中心の音声に変わったときのように、音声帯域の出入りがあったような感覚）が発生するのを抑制することができる。

【0061】このような本実施の形態 1 によるオーディオ信号符号化装置、及び復号化装置では、複数の各段の符号化器の量子化するオーディオ信号の周波数帯域を決定する特性判定部と、上記特性判定部で決定された周波数帯域と、周波数変換された元々のオーディオ信号とをその入力とし、上記複数の各段の符号化器の接続順を決定し、符号化器の量子化帯域、及び接続順を符号列に変換する符号化帯域制御部とを備え、適応的にスケラブルコーディングを行なう構成としたので、多種多様なオーディオ信号の符号化を行なう際にも、高品質、高効率の、十分な性能を発揮できる適応スケラブルコーディングを行なうオーディオ信号符号化装置、及びこれを復号する復号化装置を得ることができる。

【0062】（実施の形態 2）図 14 に、本発明の実施の形態 2 について、図 14 ないし図 20 を用いて説明する。図 14 は、本発明の実施の形態 2 による、適応スケラブルコーディングを行なう符号化装置 2001、及び復号化装置 2002 のブロック図を示す。図に示すように、符号化装置 2001 において、200105 は、符号化器の個数、ビットレート、入力オーディオ信号のサンプリング周波数、各符号化器の符号化帯域情報、等の符号化条件、200107 は複数の各段の各符号化器の量子化するオーディオ信号の周波数帯域を決定する特性判定部、200109 は符号化帯域配置情報、200110 は特性判定部 200107 で決定された周波数帯域と、周波数変換されたオーディオ入力信号とを入力とし、上記複数の各段の符号化器の量子化帯域、及び接続順を符号列に変換する符号化帯域制御部、200111 は符号化列、200112 は伝送符号化列合成器である。

【0063】また、復号化装置 2002 において、200150 は伝送符号化列分解器、200151 は符号化列、200153b は符号化列 200151 を入力とし、これを復号化する各復号化器の復号化帯域を制御する復号化帯域制御部、200154b は復号化スペクトルである。

【0064】本発明の実施の形態 2 による符号化装置 2001 は、上記実施の形態 1 と同じく、適応スケラブルコーディングを行なうものであるが、実施の形態 1 に比し、新たに、符号化装置 2001 に、復号化帯域制御部 200153 を含む符号化帯域制御部 200110 を、復号化装置 2002 に、上記復号化帯域制御部 200

0153と同じ処理を行なう復号化帯域制御部200153bを追加し、さらに、本実施の形態2の特性判定部200107においては、上記実施の形態1における特性判定部506のスペクトルパワー計算部803に代えて、図16に示すように、聴覚心理モデル計算部200602を設け、さらに、該特性判定部200107内に、符号化条件200105と、符号化帯域算出部200601より計算される符号化帯域情報200702と、配置決定部200603より出力される帯域番号200606とより、符号化帯域配置情報200109を生成する符号化帯域配置情報生成手段200604を設けたものである。

【0065】また、復号化装置2002において、200150は伝送符号化列分解器、200151は符号化列、200153bは符号化列200151を入力とし、これを復号化する各復号化器の復号化帯域を制御する復号化帯域制御部、200154bは復号化スペクトルである。

【0066】次に、本実施の形態2の動作について説明する。本実施の形態2において、符号化しようとする原オーディオ信号501は、上記実施の形態1と同様、時間的に連続するデジタル信号系列であるとする。まず、上記実施の形態1と同様な処理によって、原オーディオ信号のスペクトル505を得る。本実施の形態2では、符号化装置2001に対して、符号化器数、ビットレート、入力オーディオ信号のサンプリング周波数、各符号器の符号化帯域情報、を含む符号化条件200105を、該符号化装置2001における特性判定部200107に入力する。特性判定部200107は、複数の各段の各符号化器の量子化帯域、個数、及び接続順の情報を含む符号化帯域配置情報200109を出力し、これを符号化帯域制御部200110へ入力する。符号化帯域制御部200110には、図17に示されるように、符号化帯域配置情報200109以外に、原オーディオ信号のスペクトル505が入力され、これらを基に該符号化帯域制御部200110により制御する各符号化器で符号化を行った符号化列200111を出力し、これは伝送符号化列合成器200112へ入力されてこれにより合成され、その合成された出力が、さらに復号化装置2002へと送信される。

【0067】復号化装置2002では、符号化装置2001の伝送符号化列合成器200112の出力を、伝送

符号化列分解器200150で受け取り、符号化列200151と解析長符号化列200152とに分解する。符号化列200151は、復号化帯域制御部200153bへと入力され、該復号化帯域制御部により制御される各復号化器で復号化された復号化スペクトル200154bを得る。そして、該復号化スペクトル200154bと、上記伝送符号化列分解器200150の出力である解析長符号化列200152とから、上記実施の形態1と同様に、周波数時間変換部5、窓掛け部6、及びフレーム重ね合わせ部7を用いて、復号信号8を得る。

【0068】次に、特性判定部200107の動作を、図15～図20を用いて説明する。該特性判定部200107は、符号化条件200105を用いて符号化帯域配置情報200702を算出する符号化帯域算出部200601、原オーディオ信号のスペクトル505、及び差分スペクトル200108などのスペクトル情報、及び符号化帯域情報200702から、人間の聴覚心理モデルに基づいて聴覚重み200605を算出する聴覚心理モデル計算部200602、解析長503を参照して、これに応じて聴覚重み200605にさらに重み付けを行い、各符号化器の帯域の配置を決定して帯域番号200606を出力する配置決定部200603、及び符号化条件200105と、符号化帯域算出部200601より計算される符号化帯域情報200702と、配置決定部200603より出力される帯域番号200606とより、符号化帯域配置情報200109を生成する符号化帯域配置情報生成手段200604から構成される。

【0069】符号化帯域算出部200601は、符号化装置2001が動作を開始する前に設定する符号化条件200105を用いて、図15に示される符号化器2003が符号化する符号化帯域の上限  $f_{pu}(k)$ 、下限  $f_{pl}(k)$  を算出し、符号化帯域情報200702として、符号化帯域配置情報生成手段200604に送られる。ここで、 $k$ は符号化帯域を扱うための数で、 $k$ が0から予め設定された最大数である  $p_{max}$  になるに従って、周波数が大きな帯域を示している。 $p_{max}$  の一例は、4である。符号化帯域算出部200601の動作の一例を、表2に示す。

【0070】

【表2】

符号化条件：サンプリング周波数が48kHz、合計ビットレートが24kbpsの時

帯域 k	fpu (k)	fpl (k)
0	221	0
1	318	222
2	415	319
3	512	416

符号化条件：サンプリング周波数が24kHz、合計ビットレートが24kbpsの時

帯域 k	fpu (k)	fpl (k)
0	443	0
1	637	444
2	831	638
3	1024	832

聴覚心理モデル計算部200602は、フィルタ701からの出力信号、または符号化帯域制御部200110の出力である差分スペクトル200108、などのスペクトル情報、及び、符号化帯域算出部200601の出力である符号化帯域情報200702から、人間の聴覚心理モデルに基づいて、聴覚重み200605を算出する。該聴覚重み200605は聴覚上重要な帯域が大きな値で、聴覚上それほど重要でない帯域が小さな値となるようなものである。聴覚心理モデル計算部200602の一例としては、入力スペクトルのパワーを計算する方法を用いるものがある。入力されるスペクトルを  $x_{602}(i)$  としたときに、聴覚重み  $w_{psy}(k)$  は、

【0071】

【数18】

$$w_{psy}(k) = \sum_{i=f_{pl}(k)}^{f_{pu}(k)} \left\{ x_{602}(i)^2 * \frac{1}{f_{pu}(k) - f_{pl}(k)} \right\}$$

となる。こうして算出された聴覚重み200605は、配置決定部200603に入力され、該配置決定部200603では、解析長503を参照しながら、解析長503が小、たとえば128の時には、帯域番号200606が大である、たとえば、4、の帯域の聴覚重み200605が大きくなるように、たとえば、この帯域番号が4の帯域の聴覚重みを2倍に重み付けし、また、解析長503が小でないときには、聴覚重み200605をそのままとして、該聴覚重み200603が最大となる帯域を計算し、その帯域番号200606を、符号化帯域配置情報生成手段200604に送る。

【0072】符号化帯域配置情報生成手段200604は、上記符号化帯域情報200702、及び帯域番号200606、さらには符号化条件200105、を入力として、符号化帯域配置情報200109を出力するものである。即ち、該符号化帯域配置情報生成手段200604は、符号化条件200105を常に参照しながら、該符号化条件からして、符号化帯域配置情報200109が必要とされる間は、上記符号化帯域情報200702と帯域番号200606とを連結してなる符号化

帯域配置情報200109を出力し、これが必要で無くなるとその出力を止める動作をする。たとえば、符号化条件200105で指定された符号化器数になるまで、帯域番号200606を出力する。なお、上記配置決定部200603において、解析長503が小なときには、出力する帯域番号200606を固定する場合もある。

【0073】次に、図17を用いて、符号化帯域制御部200110の動作について説明する。符号化帯域制御部200110は、上記特性判定部200107からの出力である符号化帯域配置情報200109、および原オーディオ信号のスペクトル505を入力とし、符号化列200111、及び差分スペクトル200108をその出力とし、その内部には、符号化帯域配置情報200109を受け、原オーディオ信号のスペクトル505、及び、過去の該原オーディオ信号のスペクトル505と、該スペクトル505を符号化しかつ復号化したスペクトル200705との差分スペクトル200108を、帯域番号200606の帯域にシフトするスペクトルシフト手段200701、符号化器2003、上記原オーディオ信号のスペクトル505と復号化スペクトル200705との差分をとる差分計算手段200703、差分スペクトル保持手段200704、及び、符号化列200111を復号化器2004で復号した合成スペクトル2001001を、符号化帯域配置情報200702に基づき、スペクトルシフトを行い、これを順次合成して合成スペクトルを得、復号化スペクトル2007056を算出する復号化帯域制御部200153を含んでいる。スペクトルシフト手段200701の構成は、図20に示すとおりであるが、入力としては、シフトしたい元スペクトル2001101と、符号化帯域配置情報200109とを用いる。符号化帯域制御部200110におけるスペクトルシフト手段200701の入力のうち、シフトしたいスペクトル2001101は、原オーディオ信号のスペクトル505、または差分スペクトル200108であり、それらを帯域番号200606の帯域にシフトして、シフトされたスペクトル200

1102と、符号化帯域配置情報200109のうちの符号化帯域情報200702とを出力する。帯域番号200606に対応する帯域は、符号化帯域情報200702の $f_{pl}(k)$ 、及び $f_{pu}(k)$ から求めることができる。シフトする手順は、上記 $f_{pl}(k)$ と $f_{pu}(k)$ との間のスペクトルを、符号化器2003の処理できる帯域まで移動することである。

【0074】こうして、シフトされたスペクトル2001102を入力とする符号化器2003は、図15に示すように、正規化符号列303、及び残差符号列304を出力し、それらと、スペクトルシフト手段200701の出力である符号化帯域情報200702とを合わせたものが、符号列200111として、伝送符号化合成器200112、及び復号化帯域制御部200153へと送られる。

【0075】上記符号化器2003の出力である上記符号化列200111は、該符号化帯域制御部200110内にある復号化帯域制御部200153へと入力される。該復号化帯域制御部200153は、復号化装置2002内に存在するもの(200153b)と、動作は同じである。

【0076】次に、上記復号化装置2002内に存在する復号化帯域制御部200153bの構成を、図19に示す。復号化帯域制御部200153bは、伝送符号化列分解器200150からの符号列200111を入力として、復号化スペクトル200705bを出力するので、その内部には、復号化器2004、スペクトルシフト手段200701、復号化スペクトル算出部2001003を持つ。

【0077】上記復号化器2004の構成を、図18に示す。復号化器2004は、逆量子化部1101と逆正規化部1102とから構成されており、逆量子化部1101は、符号列200111のうち残差符号列304を入力として、該残差符号列304をコードインデックスに変換し、符号化器2003で用いたコードブックを参照し、そのコードを再生する。再生されたコードは、逆正規化部1102に送られ、符号列200111内の正規化符号列303から再生された正規化係数列303aと乗算され、合成スペクトル2001001を得る。該合成スペクトル2001001は、スペクトルシフト手段200701に入力される。

【0078】なお、符号化帯域制御部200110内の復号化帯域制御部200153の出力は、復号化スペクトル200705となっているが、これは、復号化装置2002内の復号化帯域制御部200153bの出力である復号化スペクトル200705bと同じものである。

【0079】復号化器2004によって合成された合成スペクトル2001001は、スペクトルシフト手段200701によりシフトされて、シフトされた合成スペ

クトル2001002が得られ、これは復号化スペクトル算出部2001003に入力される。

【0080】復号化スペクトル算出部2001003内では、入力された合成スペクトルを保持しており、保持しているスペクトルと、最新の合成スペクトルとを加算し、復号化スペクトル200705bとして出力する動作をする。

【0081】符号化帯域制御部200110内の差分計算手段200703は、原オーディオ信号のスペクトル505と、復号化スペクトル200705との差分を計算して、差分スペクトル200108を出力し、これは特性判定部200107へとフィードバックされる。また同時に、上記差分スペクトル200108は、差分スペクトル保持手段200704により保持されて、スペクトルシフト手段200701へも送られ、次の符号化帯域配置情報200109が入力されるときに備えるように構成されている。特性判定部200107では、符号化条件を参照しながら、該符号化条件を満たすまで符号化帯域配置情報200109を出力しつつ、それが無くなった段階で、符号化帯域制御部200110の動作も停止する。なお、上記符号化帯域制御部200110は、差分スペクトル200108を計算するために、差分スペクトル保持手段200704を持っている。これは、差分スペクトルを保持するために必要な記憶領域で、たとえば、2048個の数を記憶できるような配列である。

【0082】以上のように、符号化条件200105を満たすように、特性判定部200107と、それに続く符号化帯域制御部200110とによる処理が繰り返され、逐次、符号化列200111が出力され、それが伝送符号化列合成器200112へと送られ、解析長符号列510とともに、伝送符号化列として合成され、復号化装置2002へと伝送される。

【0083】復号化装置2002では、符号化装置2001より伝送されてきた伝送符号化列を、伝送符号化列分解器200150にて、符号化列200151と、解析長符号列200152とに分解する。該符号化列200151と、解析長符号列200152とは、符号化装置2001内の符号化列200111、及び解析長符号列510と同じものである。

【0084】分解された符号化列200151は復号化帯域制御部200153bにおいて復号化スペクトル200154bに変換され、該復号化スペクトル200154bは、解析長符号列200152の情報をを用いて、周波数時間変換部5、窓掛け部6、及びフレーム重ね合わせ部7にて、時間領域の信号に変換され、それが復号化信号8となる。

【0085】このように本実施の形態2によるオーディオ信号符号化装置、復号化装置によれば、上記実施の形態1のように、複数の各段の符号化器の量子化するオー

ディオ信号の周波数帯域を決定する特性判定部と、上記特性判定部で決定された周波数帯域と、周波数変換された元々のオーディオ信号とをその入力とし、上記複数の各段の符号化器の接続順を決定し、符号化器の量子化帯域、及び接続順を符号列に変換する符号化帯域制御部とを備え、適応的にスケラブルコーディングを行なう構成において、符号化装置には復号化帯域制御部を含む符号化帯域制御部を、復号化装置には復号化帯域制御部を設けるとともに、さらに、特性判定部におけるスペクトルパワー計算部を、聴覚心理モデル計算部とし、さらに、該特性判定部において、符号化帯域配置情報生成手段を設けた構成としたので、特性判定部のスペクトルパワー計算部にかえて代えて聴覚心理モデル計算部を用いたことにより、聴覚的に重要な部分の判定を精度よく行って、その帯域をより選択することができる。また、本発明が対象とするオーディオ信号符号化装置、復号化装置では、符号化器の配置を決定する演算を行なっている際に、符号化条件が満たされれば、符号化の処理がOKと判定されて符号化帯域配置情報も出ないこととなるが、この符号化器の配置を決定するための演算において、上記実施の形態1では、符号化器を配置するときの帯域を選択するときの各帯域幅、及び各帯域の重みが固定であるのに対し、本実施の形態2では、特性判定部の判定条件として、入力信号のサンプリング周波数と、圧縮率、即ち、符号化のビットレート、も入っていることから、これらに応じて、上記各符号化器の帯域配置を選択するときの各帯域に対する重み付け度合いを変えられるものであり、さらに、特性判定部の判定条件として、圧縮率の条件も入っていることにより、圧縮率が高いとき、即ち、ビットレートが低いときには上記各符号化器の帯域配置を選択するときの各帯域の重み付け度合いをあまり変化させないようにし、一方、圧縮率が低いとき、即ち、ビットレートが高いときには、効率をより追求するために、上記各符号化器の帯域配置を選択するときの各帯域の重み付け度合いを、聴覚上、より大事なところを強調するようにし、これにより、圧縮率と品質とのベストバランスを得ることができるものである。このように、多種多様なオーディオ信号の符号化を行なう際にも、十分な性能を発揮して、高品質、高効率な、適応スケラブルコーディングを行なう、オーディオ信号符号化、復号化装置を得ることができる。

【0086】

【発明の効果】以上のように、本発明にかかるオーディオ信号符号化方法、及びオーディオ信号復号化方法によれば、符号化ステップは、複数の符号化サブステップを有し、符号化帯域制御ステップの制御によりオーディオ信号の多段符号化を行い符号化情報を出力し、特性判定ステップは、入力されるオーディオ信号を判定し、符号化する各周波数帯域の重み付けを示す帯域重み情報を出力し、符号化帯域制御ステップは、帯域重み情報に基づ

いて、多段符号化を構成する各符号化サブステップの量子化帯域、接続順を決定し、決定した各符号化サブステップの量子化帯域、接続順に基づいてスケラブルに構成される多段符号化を符号化ステップに行わせ、決定した各符号化サブステップの量子化帯域、接続順を示す帯域制御符号列を出力するようにしたことにより、多種多様な性質を持つオーディオ信号に対して、より高音質で、より高効率な、適応スケラブルコーディングを行なうことができるという、有利な効果が得られる。

【図面の簡単な説明】

【図1】本発明の実施の形態1によるオーディオ信号符号化装置における適応スケラブルコーディングのブロック図

【図2】上記実施の形態1の符号化装置における時間周波数変換部を示す図

【図3】上記実施の形態1の符号化装置における符号化器を示す図

【図4】上記実施の形態1の符号化装置における正規化部を示す図

【図5】上記実施の形態1の符号化装置における周波数概形正規化部を示す図

【図6】上記実施の形態1の符号化装置における特性判定部を示す図

【図7】上記実施の形態1の符号化装置における符号化帯域制御部を示す図

【図8】上記実施の形態1の符号化装置における量子化部を示す図

【図9】上記実施の形態1の符号化装置における復号化器を示す図

【図10】一般のTwinVQ方式の概要を示す図

【図11】一般のTwinVQスケラブルコーディング方式を示す図

【図12】一般の固定スケラブルコーディングの短所を示す図

【図13】一般の適応スケラブルコーディングの長所を示す図

【図14】本発明の実施の形態2によるオーディオ信号符号化装置における適応スケラブルコーディングのブロック図

【図15】上記実施の形態2の符号化装置における符号化器を示す図

【図16】上記実施の形態2の符号化装置における特性判定部を示す図

【図17】上記実施の形態2の符号化装置における符号化帯域制御部を示す図

【図18】上記実施の形態2の符号化装置における復号化器を示す図

【図19】上記実施の形態2の符号化装置における復号化帯域制御部を示す図

【図20】上記実施の形態2の符号化装置におけるスペ



クトルシフト手段を示す図

【符号の説明】

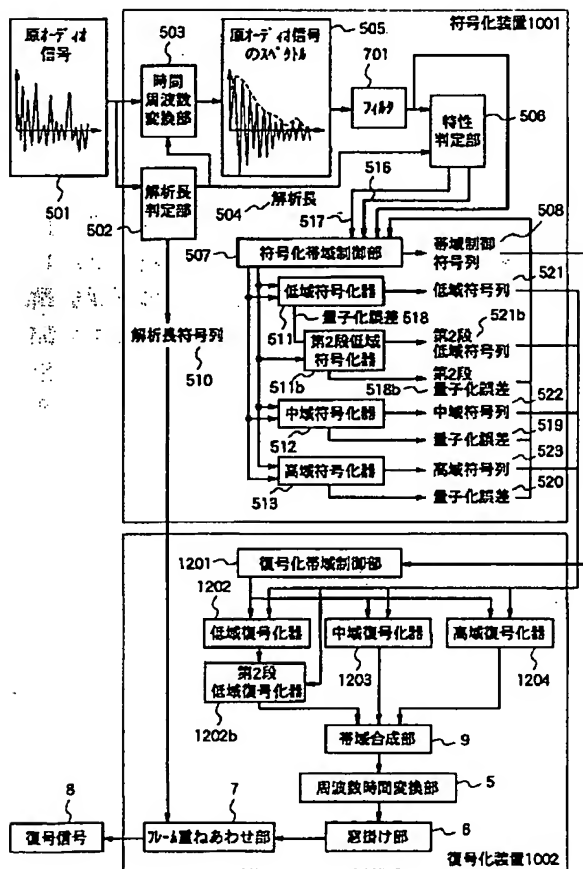
- |                     |                |        |                        |
|---------------------|----------------|--------|------------------------|
| 1                   | 符号化装置          | 516    | 符号化帯域配置情報              |
| 2                   | 復号化装置          | 901    | 帯域幅算出部                 |
| 501                 | 原オーディオ信号       | 902    | 量子化順序決定部               |
| 502                 | 解析長判定部         | 903    | 符号化器数決定部               |
| 503                 | 時間周波数変換部       | 1001   | 量子化部の量子化する帯域のMDCT      |
| 504                 | 解析長            | 1002   | 同じ量子化帯域の正規化成分          |
| 505                 | 原オーディオ信号のスペクトル | 1003   | 音源サブベクトル               |
| 506                 | 特性判定部          | 1004   | 重みサブベクトル               |
| 507                 | 符号化帯域制御部       | 1005   | ベクトル量子化器               |
| 508                 | 帯域制御符号列        | 1006   | 距離計算手段                 |
| 510                 | 解析長符号列         | 1007   | コード決定手段                |
| 511                 | 低域符号化器         | 1008   | 残差生成手段                 |
| 512                 | 中域符号化器         | 1009   | コードブック                 |
| 513                 | 高域符号化器         | 1010   | 残差サブベクトル               |
| 511b                | 第2段低域符号化器      | 1011   | ある量子化部の量子化する帯域のMDCTの残差 |
| 518, 519, 520, 518b | 量子化誤差          | 101    | 原オーディオ信号               |
| 521                 | 低域符号列          | 102    | 解析長判定部                 |
| 522                 | 中域符号列          | 103    | 時間周波数変換部               |
| 523                 | 高域符号列          | 104    | 周波数領域の原オーディオ信号         |
| 521b                | 第2段低域符号列       | 105    | 周波数概形                  |
| 701                 | フィルタ           | 106    | 正規化処理部                 |
| 5                   | 周波数時間変換部       | 107    | 正規化符号列                 |
| 6                   | 窓掛け部           | 108    | 正規化処理後の現オーディオ信号        |
| 7                   | フレーム重ねあわせ部     | 109    | ベクトル量子化部               |
| 8                   | 復号信号           | 110    | 符号列                    |
| 9                   | 帯域合成部          | 111    | 解析長符号列                 |
| 1201                | 復号化帯域制御部       | 1301   | 原オーディオ信号               |
| 1202                | 低域復号化器         | 1302   | 時間周波数変換部               |
| 1203                | 中域復号化器         | 1303   | 解析長判定部                 |
| 1204                | 高域復号化器         | 1304   | 周波数領域の原オーディオ信号         |
| 1202b               | 第2段低域復号化器      | 1305   | 低域符号化器                 |
| 201                 | フレーム分割部        | 1306   | 量子化誤差                  |
| 202                 | 窓掛け部           | 1307   | 中域符号化器                 |
| 203                 | MDCT部          | 1308   | 量子化誤差                  |
| 3                   | 符号化器           | 1309   | 高域符号化器                 |
| 301                 | 正規化部           | 1310   | 量子化誤差                  |
| 302                 | 量子化部           | 1311   | 低域符号列                  |
| 303                 | 正規化符号列         | 1312   | 中域符号列                  |
| 304                 | 符号列            | 1313   | 高域符号列                  |
| 401                 | 周波数概形正規化部      | 1314   | 解析長符号列                 |
| 402                 | 帯域振幅正規化部       | 2001   | 符号化装置                  |
| 403                 | 帯域テーブル         | 2002   | 復号化装置                  |
| 601                 | 線形予測分析部        | 200105 | 符号化条件                  |
| 602                 | 概形量子化部         | 200107 | 特性判定部                  |
| 603                 | 包絡特性正規化部       | 200108 | 差分スペクトル                |
| 803                 | スペクトルパワー計算部    | 200109 | 符号化帯域配置情報              |
| 804                 | 配置決定部          | 200110 | 符号化帯域制御部               |
| 517                 | 帯域制御重み         | 200111 | 符号化列                   |
|                     |                | 200112 | 伝送符号化列合成器              |



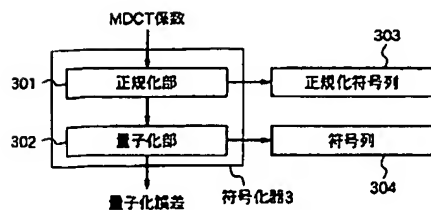
200150 伝送符号化列分解器  
 200151 符号化列  
 200152 解析長符号化列  
 200153 復号化帯域制御部  
 200154 復号化スペクトル  
 2003 符号化器  
 200305 符号化帯域情報  
 200601 符号化帯域算出部  
 200602 聴覚心理モデル計算部  
 200603 配置決定部  
 200604 符号化帯域配置情報生成手段  
 200605 聴覚重み

200701 スペクトルシフト手段  
 200702 符号化帯域情報  
 200703 差分計算手段  
 200704 差分スペクトル保持手段  
 2004 復号化器  
 200901 逆量化部  
 200902 逆正規化部  
 2001001 合成スペクトル  
 2001002 シフトされた合成スペクトル  
 2001003 復号化スペクトル算出部  
 2001101 元スペクトル  
 2001102 シフトされたスペクトル

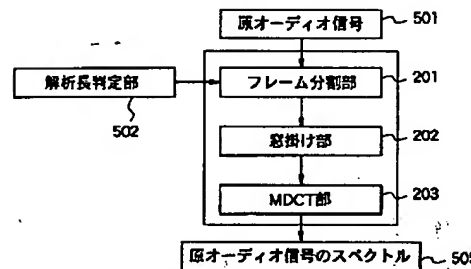
【図1】



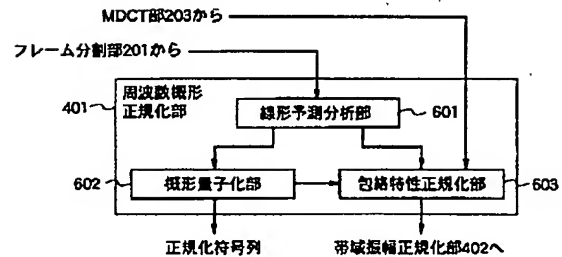
【図3】



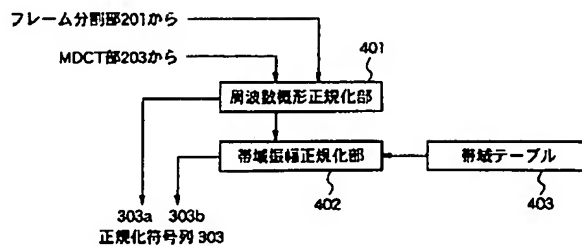
【図2】



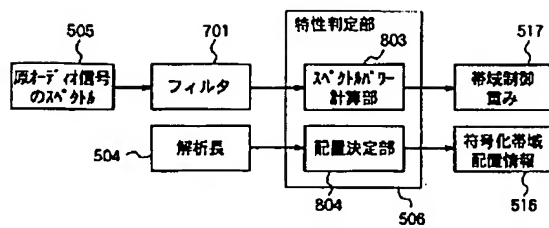
【図5】



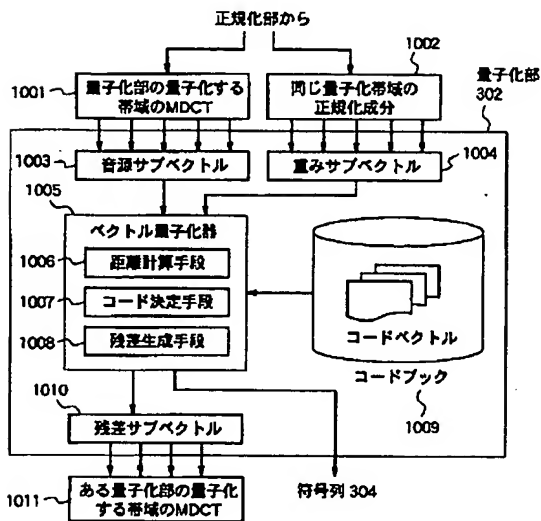
【図4】



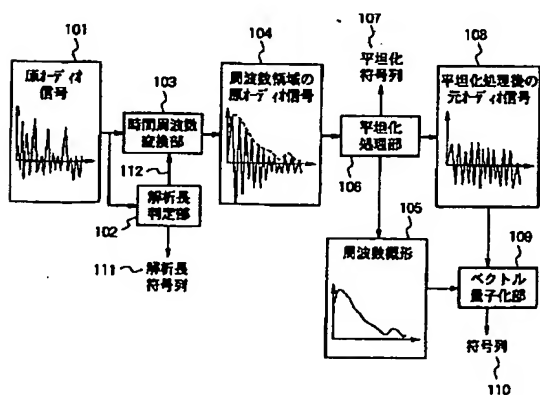
【図6】



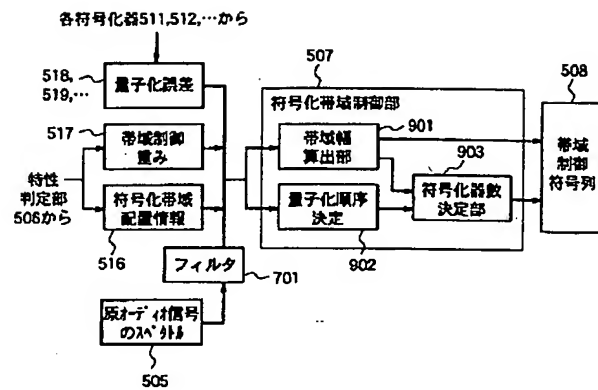
【図8】



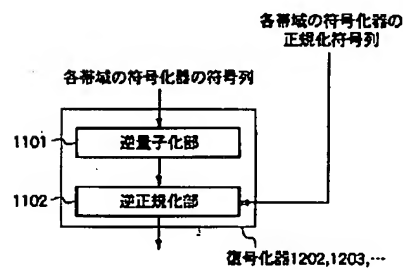
【図10】



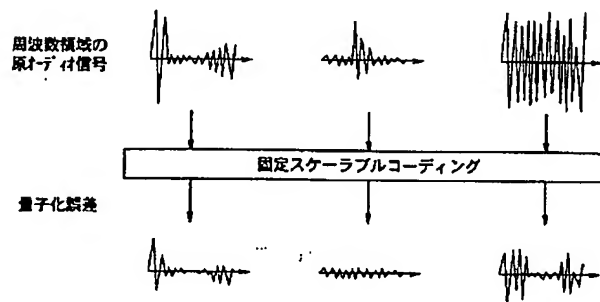
【図7】



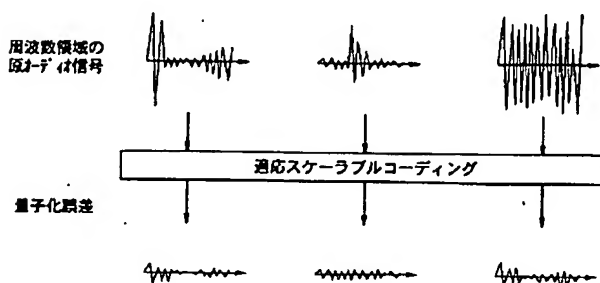
【図9】



【図12】

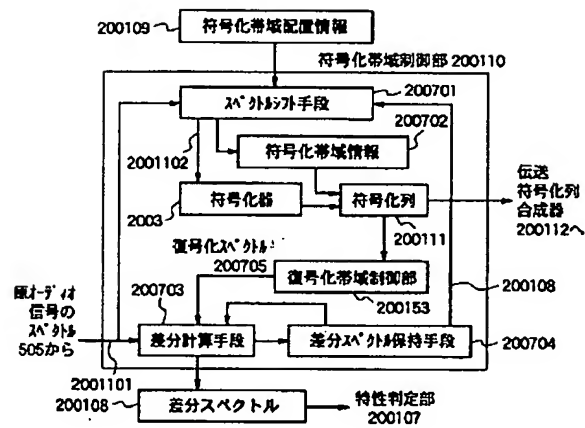


【図13】

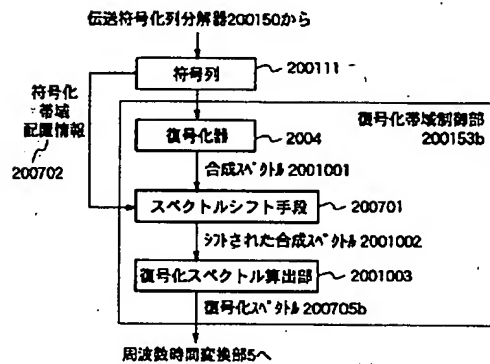




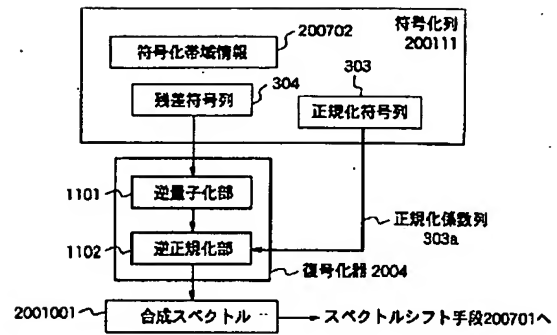
【図17】



【図19】



【図18】



フロントページの続き

(72)発明者 石川 智一

大阪府門真市大字門真1006番地 松下電器  
産業株式会社内

Fターム(参考) 5D045 DA20

5J064 AA01 AA02 BA04 BA13 BA16  
BB12 BC01 BC02 BC11 BC14  
BC16 BC17 BC18 BC26 BD01